**COMPETITIVENESS AND INNOVATION FRAMEWORK PROGRAMME
ICT Policy Support Programme (ICT PSP)**

e-CODEX

e-Justice Communication via Online Data Exchange

D6.1 Requirements

Abstract:

For e-CODEX implementing electronic services for cross border legal and judicial procedures Work package 6 (WP6) is responsible for document standards and semantics. WP6 defines a methodology and develops specifications to enable meaningful exchange of information between electronic systems supporting legal procedures within a supra national context.

The European Interoperability Framework presents requirements, principles and guidelines that are a foundation for the semantic interoperability layer and other layers. This deliverable refines and adds to these guidelines to meet the legal and judicial perspective. The basic principles of the EU as subsidiarity and creating a level playing field for the single market are respected.

History

Version	Date	Changes made	Modified by
0.5	29 April 2011	First draft	NL
0.6	2 June 2011	Remarks quality manager	NL
0.7	4 June 2011	Remarks MS + update requirements	NL
0.9	7 June 2011	Update according to comments from internal review	NL
1.0draft	15 June 2011	Update according to comments from MS + remarks internal quality manager + added appendix requirements	NL
1.0 final	30 June 2011	Update according to comments from MS + external quality manager	NL

Table of contents

List of figures

List of abbreviations

Acronym	Explanation
CIP	Competitiveness and Innovation Programme
D6.1	Deliverable 6.1:requirements
D6.2	Deliverable 6.2: List of Standards, reusable Assets to be used and missing building blocks'
e-CODEX	e-Justice Communication via Online Data Exchange
ECRIS	Exchange of information extracted from criminal records
EIF	European Interoperability Framework
epSOS	European Patients Smart Open Services
EQM	External Quality Manager
EU	European Union
Eurojust	A judicial cooperation body created to help provide safety within an area of freedom, security and justice
EU MS	European Union Member States
US NIEM framework	United States National Information Exchange Model framework
e-Government	Electronic Government
e-Services	Electronics Services
e-Signatures	Electronic Signatures
IDABC	Interoperable Delivery of European e-Government Services to public Administrations, Businesses and Citizens
ICT	Information and Communication Technology
ICT PSP	ICT Policy Support Programme
L12B	Legal institution to Business
L12C	Legal institution to Citizen
LI2LI	Legal institution to Legal institution
LSP	Large Scale Pilot
OCD	Omnivarious Company Dossier
ODF	Open Document Format
OWL	Web Ontology Language

PDF	Portable Document Format
PEPPOL	Pan-European Public Procurement Online
RC	Recommendation
RFC	Request for Comments
RQ	Requirement
SEMIC	Semantic Interoperability Centre Europe
SPOCS	Simple Procedures Online for Cross- Border Services
STORK	Secure idenTity acrOss boRders linKed
URI	Unique Reference Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
URC	Uniform Resource Characteristics
UTF-8 format	8-bit Unicode Transformation Format
VCD	Virtual Company Dossier
W3C	World Wide Web Consortium
WP	Work Package
XHTML	eXtensible HyperText Markup Language
XSLT	eXtensible Stylesheet Language Transformations
XML	eXxtensible Markup Language
ZIP file	A data compression and archive format

Prologue

ONE day I wrote her name upon the strand,
But came the waves and washèd it away:
Again I wrote it with a second hand,
But came the tide and made my pains his prey.
Vain man (said she) that dost in vain assay
A mortal thing so to immortalise;
For I myself shall like to this decay,
And eke my name be wipèd out likewise.
Not so (quod I); let baser things devise
To die in dust, but you shall live by fame;
My verse your virtues rare shall eternise,
And in the heavens write your glorious name:
  Where, when as Death shall all the world subdue,
  Our love shall live, and later life renew.

Edmund Spenser

One day I wrote her name upon the strand,[1]…
To be semantically interoperable we define a methodology, an inter lingua and standardize. A good solution stands the test of time. The poet Edward Spenser[2] and his lovely lady are immortal through a sensitive poem. We must serve the European Union. A community that is multi cultural by nature: different nations, different customs, many languages and different legal systems.

A single market in the European Union requires cross border information exchange, in order for communication to flow endlessly in waves…

The European Union grows and gives opportunities to enterprise and flourish.
Electronic communication is the modern way we collaborate and join forces. In poetry trust, truth and love are small words that don't require precise explanation.
However, for real live collaboration we have to fine tune and coordinate our efforts, because legal and judicial procedures have to be precise, sound and clear.

The quest we start with a first step…, Functional requirements.

---

[1] The first sentence from the sonnet 75 Edmund Spenser wrote.

[2] Edmund Spenser (c. 1552 – 13 January 1599) was an English poet best known for The Faerie Queene, an epic poem and fantastical allegory celebrating the Tudor dynasty and Elizabeth I. He is recognized as one of the premier craftsmen of Modern English verse in its infancy, and one of the greatest poets in the English language. Source Wikipedia: http://en.wikipedia.org/wiki/Edmund_Spenser

Executive summary

e-CODEX stands for e-Justice Communication via Online Data Exchange. e-CODEX is the fifth pilot A from the ICT Policy Support Program, with predecessors like epSOS, STORK, PEPPOL and SPOCS. The goal of e-CODEX is to **improve the cross-border access for citizens and businesses to legal means** in Europe as well as to **improve the interoperability between legal authorities** within the EU. Document standards and semantics, WP6, defines specifications to enable meaningful exchange of information between the systems supporting legal procedures across the Member States by enabling semantic interoperability.

The European Union evolves and requires different forms of communication to cope with the growth in data exchange caused by increasing commercial activities. Exclusively manual processing does not give the responsiveness a modern society requires. With e-CODEX automation for legal procedures with support of e-Services, system-to-system integration and message transport becomes more mature.

To improve interoperability, European nations first have to be able to interpret information to be sent and received. To share information, the reason why and the semantics have to be clear to all communication partners. Semantic components that represent the information to be shared have to be in place. Initiatives to ensure just this are growing in the European Union.

WP6, defines a methodology and develops specifications to enable meaningful exchange of information between electronic systems supporting legal procedures within a supra national context. WP6's responsibility and focus is on the semantic interoperability layer[3]. WP6 joins its effort with existing initiatives, covering other domains, to introduce general and business case specific semantic components.

Types of interaction to be supported in e-CODEX are LI2B, LI2C and LI2LI. These interaction types are formed by four types of communication: publication, forms, intelligent forms and system-to-system. These four methods to communicate, their application and different levels of complexity result in different options of choice for use cases to implement.

e-CODEX prioritizes system to system integration. WP6 therefore develops a methodology to structure messages ready for machine interpretation. WP6 concentrates exclusively on the message design. WP5 develops instruments for routing the message and WP4 is responsible for signing the message.

Legislation leads to legal and judicial procedures which in turn lead to business processes supporting the execution of the legal and judicial procedures. The business process level is where the information is exchanged when executing cross border legal and judicial procedures. WP6 places the

---

[3] The third of the four interoperability layers presented in the European Interoperability Framework version 16 December 2010

legal and judicial procedures and the business processes in the middle and covers them with WP6's methodology to facilitate cross border legal activities.

For European nations to be semantically interoperable they have to share a technical language, its vocabulary, syntax and semantics must be well understood. With a common technical language communication partners compose messages that are easy to process and are consistent. For system-to-system integration and electronic message exchange additional standardisation is required.

To model data and messages WP6 identifies an architectural framework consisting three individual layers: the conceptual, logical and technical layer. The conceptual model and its logical equivalent are input for the technical level, the implementation. To work with this three stage approach where every stage builds on the previous one provides a robust methodology.

Each stage may have its own perspective and strategy to reach result. Conceptual modelling is the first stage where data elements are identified and put into context. The conceptual model specifies data elements and may incorporate additional criteria to be met. The second stage is to create and fine tune generic building blocks for electronic messaging and to define possible transformations for these building blocks and messages. The third stage involves a technical specific implementation and generating the technical message specification.

WP6 proposes a methodology to support a new level of semantic interoperability where a clear interpretation for information is given. System to system integration may evolve in close relation with legacy systems present by decoupling the technique and the logical specification. Combining and transforming semantic building blocks is made possible obeying supra national and national criteria.

The European Interoperability Framework addresses core requirements that form the foundation for the requirements presented in this document.

The general requirements presented here need fine tuning for each of the use cases selected in WP3, pilots.

A first impression of the EU initiatives PEPPOL[4], SPOCS[5], Eurojust[6] and SEMIC[7] offers examples for guidelines, standards and results ready for adoption are given. The knowledge and support of these e-CODEX allies give input for further development of e-CODEX.

---

[4] Pan-European Public Procurement Online   http://www.peppol.eu/

[5] Seamless cross border electronic procedures   http://www.eu-spocs.eu/

[6] Eurojust is a judicial cooperation body created to help provide safety within an area of freedom, security and justice. http://www.eurojust.europa.eu/

[7] Semantic Interoperability Center Europe   http://www.semic.eu/semic/

1.   Introduction

1.1.   Purpose

e-CODEX (e-Justice Communication via Online Data Exchange) is the fifth pilot the ICT Policy support program has initiated. The goal of e-CODEX project is to improve the cross-border access of citizens and businesses to legal means in Europe as well as to improve the interoperability between legal authorities within the EU. e-CODEX initiates improvements for cross border electronic communication supporting judicial and legal procedures. Improving the means to communicate electronically gives rise to new and better e-Services. Beside the legal institutions, both citizens and enterprises receive more and better channels to communicate in respect to judicial and legal procedures.

Work package 6, document standards and semantics (WP6) will deal with the different European cultures and systems that affect interpretation by thoroughly analyzing cross-border processes and deriving standards and mapping schemes for connecting distinctive national systems. Respecting national solutions is not only a sound approach because of the investments to build these systems. National solutions also have taken into account the peculiarities of the legal systems of each of the Member States. These peculiarities might prove to be difficult to overcome when creating a pan-European system. Assessing the processes of legal and judicial organizations of the Member States by respecting existing solutions, enlarges adoption of the e-CODEX solution.

The agreements to be made for a business interoperability interface to work have their technical counterpart. e-CODEX honours the subsidiarity principle in providing an architecture for interoperability and brokering data. This includes semantic components that are technology neutral by nature. National e-Services may continue to operate with an additional interface for European communication and services. This strategy preserves the previous investment made for national e-Services.

Interoperability, achieved by interpreting document content and metadata amongst European IT-Systems delivers a major push towards more effective cross-border processing of legal and judicial transactions. The use cases of e-CODEX demonstrate how data and semantics are processed through interoperable IT-Systems thus increasing the efficiency of cross-border processing.

WP6 concentrates on document standards and semantics. This document describes the high level requirements in order to enable a mapping of existing solutions. Input for this functional specification are the meetings and teleconferences organised by the WP6 project team.

1.2. Methodology of work

For legal and judicial European information exchange requirements have to be identified, evaluated and refined. This document incorporates requirements that are general by nature. Since the use cases have not been selected at the time of writing, these will be assessed in a later stadium to come up with use case specific requirements.

The European Interoperability Framework is the European framework for generic requirements and guidelines. The requirements from this framework are taken into account in the requirements for WP6. The specific requirements for e-CODEX WP6 should be and are in line with the requirements from EIF.

The WP7 Standards & Guidelines regarding document standard and semantics are mandatory for the WP6 requirements. The uses cases are more specific and will need context specific requirements. Each use case is unique and might use a different type of interaction which could lead to additional technical requirements. The upcoming task for WP6 is to enlist existing solutions and identify their reusable components.

Large scale pilots like PEPPOL and SPOCS present European solutions for cross border information exchange. The initiatives SEMIC and Eurojust present European solutions for vocabularies. All four of them are taken into account and are an inspiration for WP6 (see Appendix II). The responses to a questionnaire sent out in December 2010, offers a high level overview of European and national solutions for interoperability. That input is used in the meetings of WP6 leading to this document.

The requirements reflect the scope and present a three layer methodology for system to system integration to support cross border legal and judicial procedures.

2.   Scope and goals

2.1.   Global

As already mentioned the goal of e-CODEX is to improve the cross-border access of citizens and businesses to legal means in Europe as well as to improve the interoperability between legal authorities within the EU.

For European interoperability a new era starts. The European citizens are free to work, live, trade, travel and settle anywhere, any time in the EU. Enterprises operate freely in **a single market** environment. Wherever there is cross border activity, governance and legislation provide the necessary means to commercial parties to operate.

To enable both citizens and enterprises starting cross border legal and judicial procedures, **legal and judicial communication services** have to be in place. By nature the EU is a multi cultural and multi lingual society. This asks for **standardization of document standards and semantics** in order to be able to communicate understandably about legal and judicial procedures for a European level playing field.

2.2.   WP6 General

e-CODEX gives life to new technical improvements and interoperability for the legal and judicial domain. WP6 concentrates on the **semantic interoperability layer** and **system to system** integration. The other work packages concentrate on architecture, infrastructure, electronic messaging, routing, e-Identity and digital signature. In combination e-CODEX delivers a sustainable European legal and judicial communication high way.

e-CODEX works use case centric. **The European Interoperability Framework** gives requirements and recommendations to be interoperable. e-CODEX follows these guidelines[8]. The general requirements are ready for refinement for the use cases that are subjects for the pilots to come[9].

WP6 'document standards and semantics' defines methods, standards and specifies services for semantic interoperability. To be interoperable both structure and content for documents and messages have to be present.

---

[8] EIF 1: use EIF, the recommendations and guidelines

[9] EIF 4: a demand driven approach for identification of requirements

The nature of the legal and judicial services gives rise to different requirements and to the amount of structure to provide. These requirements translate into guidelines and standards how to structure the content, document, message and additional metadata. For system to system integration **exclusively structured XML messages** and **modelling data elements** are taken into account.

2.3. Three levels

Independent of the selection of use cases within the pilots the following **three levels to engage the use cases** are taken into account:
1. Legislation
2. Legal and judicial procedure
3. Business process

These layers have their own impact on semantics and might be contra intuitive in the details of implementation. The legal and judicial procedure is a derivation of the legislation. Both national and European procedures influence the information exchange in the actual business processes. WP6 will come up with a method to guide this.

2.4. Multilingualism

As the EU is a multi lingual society, providing support for cross border procedures automatically involves multi lingualism. Translation of information, automated or manually, is not part of the e-CODEX project. However, WP6 aims at adoption of solutions present for translation. Where appropriate WP6 references these solutions but no active effort is made in development of services for translation.

2.5. Modeling and design

WP6 is responsible for defining a methodology to develop data definitions and support the reuse of available data definitions. In addition WP6 is responsible for specifying naming and design rules for creating electronic XML messages and data elements. The data modelling supports the implementation of legal and judicial procedures with electronic services.

WP6 commits itself to active participation in the different use cases under development and part of the e-CODEX project. WP6 will support WP3 in identifying essential information for the business processes executing legal and judicial procedures by offering the right data elements. WP6 will support also the party found responsible for the interaction design for intelligent forms, again by identifying the essential information and by offering the right data elements representing that information.
The communication partners and technical levels of sophistication may differ for each specific use case. For each use case the relevant requirements present in this document will be taken into account.

Out of scope are the design of direct human computer interaction and the ergonomic requirements for presentation of information.

2.6. Summary

The EU aims for the single market. Governance and legalisation provide a stable foundation for this single market. The legislation and the electronic support of legal procedures stimulate cross border activity and by that further economic growth.

e-CODEX follows a use case centric approach for the development of electronic services for legal and judicial procedures. WP6 concentrates its efforts on the semantic interoperability layer to support system to system integration for legal and judicial procedures. Modelling data elements that can be combined into meaningful electronic XML messages forms the foundation.

Three levels to engage the semantic aspect of the use cases are taken into account, legislation, legal and judicial procedure and the business process. WP6 will come up with a method to overcome potential contradictions between the three levels.

WP6 references existing services for translation. WP6 excludes the development of translation services from its program as other parties already provide these services. Further exclusions are: direct interaction with the end user, the presentation format, user interface and document layout.

3. Semantic interoperability

The European Union prospers where communication is effective and efficient. Member states communicate both in bilateral and in multilateral senses. Where different partners are involved there is the need to take additional care in synchronisation and mutual agreement.

Semantic rules and constructs play a key role in making communication possible in a multilingual and multicultural setting. To communicate, partners have to agree upon the use case for information to exchange and the purpose for the individual messages but also their relation. A human interpretation depends on customs and the language of speech. The written word requires attention to the letter. Semantic rules and constructs give a simplification to achieve a practical result.

European legislation and national legislations have their own perception on legal matters. For legal and judicial procedures on a supra-national or bilateral level both perspectives have to be taken into account. Member states have to be aware of the local differences and require a general vocabulary to transform from a European perspective to the national interpretation.

Civilians and enterprises which start a legal and judicial procedure or become subject of a procedure have to be informed on how to respond and what steps are to be taken. The procedure is to be presented in a clear format and must be easy to comprehend. 'WP6 documents and standards' develop guidelines on how to structure the content part of a procedure and ready to be sent over the wire.

In close relation to each other, the e-CODEX work packages define a methodology for interoperability. The member states align where a cross border communication channel is to be realized. Infrastructures on a technical and logical level (WP5) are to be present before communication can take place. A communication protocol on how to interact gives essential information concerning what data to expect and when messages are routed on the basis of an address or content.

e-CODEX covers different legal domains. To express those domains we indentify concepts, their relations and building blocks for usage in real communication. Different models have different purposes. A conceptual model may have a counterpart down the workflow to express more details not necessary for a first specification. E.g. selections of attributes and data types are not part of the conceptual model. In addition to concepts, an intelligent computerized system requires facts and rules to operate. To streamline legal and judicial procedures the human expertise and knowledge is to be deduced. An upper level ontology, a formal conceptual model to express knowledge, and domain specific ones give support for defining criteria to be met. These criteria prescribe what type of content is part of the procedural steps.

Human interpretation is of great importance and forms the foundation for communication. Data elements are the atomic parts. Only related to each other they get a meaning. A document or message incorporates predefined data elements to enable a consistent interpretation. Sharing information and an agreed interpretation for the data to send and receive is the foundation for the semantic interoperability layer.

To automate the human interpretation there is the need of giving guidelines on how to construct, relate and transform the data. A use case gives context. Partners have to align their objectives and the process requires different iterations in order to agree on the mandatory data. To work active on multilingual issues increases the human involvement to define the appropriate data elements and map them in a way all parties understand their meaning.

4. European interoperability framework

The European Union is a society. A single market within the community gives impulses for economic growth. A society where people travel, where there is free movement of goods and cross border business activity. Therefore legal considerations have to be taken into account. More activity gives rise to better and richer set of e-Government services. The possibility for the public to get services on demand when required will increase acceptance and approval of EU policies and procedures. The European Interoperability Framework gives guidelines and requirements to ensure the quality and modularity of those services. WP6 takes the requirements as input to show the relation between e-CODEX and EIF.

4.1. The scope

For European interoperability the European commission prescribes recommendations to follow. These recommendations are part of the European interoperability framework[10]. They give guidance on how to form interoperable layers and services for sharing information. The recommendations give a good foundation for e-CODEX.

> *Requirement WP6_RQ_F1*: eCODEX SHOULD incorporate the EIF recommendations and refine them for legal and judicial application.

4.2. The major goals

The European Commission and Member States give direction towards a European society where Member States are interoperable in a non technical manner with technical support[11]. This human oriented viewpoint towards computerization is the essence to be semantically interoperable.

Member States carry responsibility for both the national and the supra-national level to be interoperable. Member States standardize and define European eGoverment services together.

The e-Services give the European society a chance to interact with the government in a way not shown before. The multicultural and multilingual nature of our European society is displayed in the way the services are implemented.

The e-Services give support for legal and judicial procedures so people are invited to use them when necessary. The human interaction stands on the fore ground. The high quality of the services ensures accessibility for every citizen. Great care is taken into making the services secure and privacy sensitive.

---

[10] EIF 1: use EIF, the recommendations and guidelines

[11] EIF 2: principles of a general nature

WP6 concentrates her efforts on system to system integration. A clear separation is made between the presentation and the data structure. To present a generic data structure for an electronic message and to provide a method to define the semantics for the individual data elements present in the electronic message are the major goals for WP6. The focus on the business interoperability interfaces between MS sets collaborations free to engage in cross border information exchange pilots.

The services are transparent and open. Open standards and open source based solutions give room for improvement by the community. They nurture multilateral platforms and solutions.

4.3. The layers of interoperability

The layers[12] of interoperability form a stack. Governance and legislation are the foundation and input for implementing business processes. These business processes implement legal and judicial procedures. For cross border information exchange definitions and structures are part of the semantic layer. The documents and messages that are formed in this layer are sent and received using the middleware and technical infrastructure, which are part of the technical architecture.

The work packages follow closely the layers of interoperability and build on each other. The pilots are a special case where all deliverables from the work packages are combined and put to practice.

WP3 Pilots implement takes all the layers into account. The methodologies for business processing and semantics defined in WP7 and WP6 are in use to define the necessary models. WP5 gives a good foundation for middleware and technical infrastructure.

WP7 Architecture gives overall guidance for the architectural principles and has a major part in quality assurance. WP7 identifies the standards and topology. WP6 incorporates these products and guidelines and refines them for the semantic interoperability layer.

Where documents require additional privacy measure, security or proven authenticity WP4 delivers the technical solutions. E.g.:' e-Identity', digital signature, watermark and other cryptographic services.

WP5 is the cornerstone for electronic interoperability. The middle ware, communication protocol and routing are the infrastructure to send and receive data. WP6 defines the methodology and principles to structure the documents and messages sent over the wire. A starting point for the methodology is the recognition that WP6 focus on the semantic interoperability layer. Even so the interaction with the other work packages and their requirements is of great importance. WP7 in conjunction with WP6 prescribe the open standards for the semantic interoperability layer. To specify structure, syntax rules, criteria and additional supporting constructs, a three stage development methodology is present. In the chapters 7 '(The Conceptual World'), 8 '(The Logical Layer') and 9 '(The Technical Implementation') further directions are given.

---

[12] EIF 3: Different layers, different viewpoints considering interoperability

WP2 Communication gives operational support to inform the public, the business and the governments.

4.4. Forms of interaction

The identities of the partners who communicate with each other are essential for selecting the e-Service and the proper functionality.

Three basic types of interaction are present.
- Government to government
- Government to business vice versa
- Government to citizen vice versa

Government to government information sharing that is purely governmental and bi-lateral in nature is distinguished from communication to resolve cases that citizens and enterprises initiate.
For government communication with the public, citizens and enterprises more than one Member State and or institution may be involved. The regional and or national solutions for cross border information exchange must appear as a single point of contact to the user.

Besides the above actors, legal professionals and judiciaries are specific groups to pay attention to.
EIF is generic and doesn't discriminate between more specific roles wherein a party operates.
LI2B, LI2C and LI2LI in combination with role specific needs are the e-Justice addition to EIF.

For all interaction forms, more than one member state or institution or agency may be involved. Furthermore, different constellations and use cases imply different strategies for communication and levels of sophistication. Four levels are present; as detailed hereby in the next subsections.

4.4.1. Level 1: Publish

Web publication concerns people visiting a web site reading information online. Publishing information on the internet allows reaching and informing the public and enterprises on a large scale. A web site possesses additional functionality to ease the process of finding the right material. Some widely accepted means are a search engine, a dictionary and linking documents. Multilingual search and translating documents to different languages are additional features. Public availability of legislation is a citizen's right. To publish the legislation on the internet is a step further. Legal and judicial procedures may reference the legislation and give an easy way to look up relevant regulations and directives. The creation of publication material is out of the scope for WP6. Constructs for referencing this material are taken into account.

4.4.2. Level 2: Form

A downloadable form or document is a very basic step in automation when it comes to inquire data. People are able to fill in the form offline and are free to send in the form when suitable for them. Both, a template or a free format are options to get the required data. Furthermore, sending in the form may be by electronic means or by traditional post. The latter is out of scope for e-CODEX. For e-CODEX web sites and email are considered suitable communication channels.

4.4.3. Level 3: Intelligent form

Filling in a form may be a tedious undertaking. A standard form requires additional text to guide a person on what to fill in. Using an intelligent form, the experience can be improved by giving additional support in giving a person step by step guidance. Filling in one field leads to displaying an additional field. The data input given by a person provides the system with essential information to decide what to display next. For defining and verifying criteria to be met for a legal procedure, the same technology gives a solution to verify an electronic message for those criteria.

Interaction and form design aren't part of WP6. The criteria to be met and the entities that may occur are subject for development in WP6. The conceptual model is the place where those definitions live.

4.4.4. Level 4: System to system

WP6 concentrates her efforts on level 4 system-to-system integration. It involves direct communication, where systems communicate without human intervention. The essence is to process data automatically when it comes available or when the information is needed to fulfil the needs of a legal procedure. Business processes have system support. A process triggers a next event where action has to be taken. Sending data to a business partner may then be an automated decision. Also where bulk processing is a major task, system-to-system communication is a logic step in automation.

# 5.	Semantic interoperability in depth

## 5.1.	Data and information

Semantic building blocks are made of data. Information is the human interpretation given to an aggregation of data. Information serves a purpose. For decision making we require information. To define data and to have a semantic correct interpretation a proposition is made to develop in a three stage modelling mode. The three stages: the conceptual model, the logical specification and the technical implementation are common for database design purpose. For communication and information sharing in a distributed sense they are less common.

The conceptual model is the foundation for the information architecture. The logical specification is a more in depth specification of the entities that form an information package. The technical implementation of documents, intelligent forms, messages and information systems depend on the logical specification. The technical implementation is technique dependent whereas the logical specification and conceptual model are independent from technical considerations.

Recommendation WP6_RC_F1: The information architecture for WP6 SHOULD be formed by a three layer methodology that identifies the conceptual model, logical model and technical implementation.

Member States shall collaborate to define common data elements[13] to be used in e-Government services. For European interoperability, data elements and their definitions have to be present, both on an European and national level This to be able to make a proper mapping: national to European or national to national.

Requirement WP6_RQ_F2: An European core model SHOULD be present for e-CODEX to link sector and use case specific models.

Requirement WP_RQ_F3: Data elements for data exchange purpose SHOULD be defined on the logical model layer.

Requirement WP_RQ_F4: Member States SHOULD map their national data models with the European core model.

---

[13] EIF 7: national, sector and core pan-European data elements plus mappings

5.2. Sharing information

e-CODEX focuses on documents as the exchange data carriers. The documents may be a document people are able to download, a web form or an electronic message.

5.2.1. Structure or structureless

Communication takes place with more or less structure. A document may be a free form format or adopt a standard structure. A document may stand on its own or belong to a dossier.

Three primary formats concerning the structure of information are present.
- Structured format
    - Both human and machine are able to process and interpret this format.
    - Tight system-to-system integration becomes possible.
- Unstructured formats
    - Human interpretation and processing is dominant for this format.
    - Additional automated enhancements have a supportive role for the human interpretation given.
- Container formats
    - This format may contain different types of data and documents.
    - The documents form a logical work package or case.

WP6 works on system to system integration where systems send and receive data in a machine interpretable XML format. The structure is conform specific naming and design rules. The XML elements are part of a dictionary.

5.2.2. Presentation

Data and structure in combination with human interpretation leads to information. The presentation form may be relevant to ease the human ability to process the data. A document with both textual and graphical representations gives the human reader more guidance than a document with textual data only. Furthermore, preserving the original format of a document may be an essential step for official reference.

> *Requirement WP6_RQ_F5:* The original format of a document SHOULD remain intact during transportation

5.2.3. Forms of distribution

WP7 Architecture defines the enterprise architecture. The star topology, communication patterns, services and their interaction are under consideration.
In addition WP7 prescribes how data is to be delivered for legal and judicial requirements.

5.3. Data storage

Data storage is a National responsibility.

> *Requirement WP6_RQ_NF1:* European and national guidelines SHOULD be leading concerning digital archiving legal cases.

5.4. Retrieve information

> *Requirement WP6_RQ_F6:* All meaningful objects which are exchanged SHOULD have unique and persistent identifiers for later reference.
>
> *Requirement WP6_RQ_F7:* Where identification systems for legal information are available they MUST be used.
>
> *Recommendation WP6_RC_F2*: The Dublin Core SHOULD be in use to describe resources.

6.   The conceptual world

Use cases for cross border information exchange and their analysis involve defining concepts. Concepts are an abstraction and simplification of real objects. Related they become a first framework for further work.

Use cases' goal and purpose require clarification in systematic way. An agreement on the interaction type and the data to exchange is a prequisite. A division in the responsibility for designing the business process and the data elements is made. WP7 architecture describes the business process; WP5 infrastructure gives requirements for system to system interaction. The business partners have to apply a common language to guarantee interoperability. WP6 is thereby responsible for identifying, defining and relating concepts, and deriving building blocks for electronic message exchange.

Information and data specialists identify the relevant information and data elements. The business process is the first step in identifying the data elements and components for communication. The concepts under consideration find their place in the conceptual model.

6.1.   A conceptual model

To clarify the world around us we form a conceptual model. A conceptual model is an abstraction of the real world. The model gives structure and relates concepts to each other. A concept is a definition for a real or abstract entity. An entity we perceive with our senses is a real concept. An abstract entity is a generalization or a formalization of binding constructs. We relate concepts to strengthen the model and the meaning. In combination the concepts, their relation and definition, all together are the model.

The conceptual model is suitable for different purposes. Identifying concepts to be in use for cross border information exchange is a first step to create a mapping for bilateral and multilateral information exchange. The real multilingual mapping takes place on the logical level. In the conceptual model, general metadata elements and terms to define criteria are best described in a language that is comprehendible and acceptable for all nations. The English language is suitable for translations of general concepts. For more in depth translations an intermediary language is not sufficient.

> *Recommendation WP6_RC_F3:* An inter lingua SHOULD be present to develop a conceptual model.

For European regulations, concepts for use in the legal and judicial procedure are to be identified. For directives, both a European model and national model are used to present the legislation. To match both perspectives certain criteria have to be met. The criteria shall be expressed in a model.

A conceptual model may incorporate and reuse other conceptual models. More general concepts, for instance natural person, may have different counter parts for each domain e.g. lawyer, citizen of state and operator. e-CODEX is use case centric. The conceptual models reflect this work modus. Every domain specifies its own set of conceptual models.

The specification of concepts is done in a technical neutral way. Functions like linking, merging, importing and exporting models are application specific.

6.2. The paradigm

A conceptualization may be modelled in different ways. The methods and the possible constructs are, in the conceptually, world part of the whole. A paradigm is a design discipline with a method and set of constructs taken from the spectrum of possible combinations. The richness of the constructs, their purpose, their expressive power and the rules to combine them characterize the paradigm.

Three paradigms are in place for the conceptual world:
1. Entity relationship modelling
2. Object oriented modelling
3. Logical formalization

The entity relation model is a well know paradigm to model databases. Entities, their attributes and their relations are present. The model is very basic. Additional constructs for modelling are given by the object oriented style, for instance inheritance expresses child parent relations.

An object oriented modelling style gives advantages above the traditional entity relation modelling style. The model feels more natural and gives more structure to add new statements. Attributes are common amongst different entities. It is easy to introduce a new sub set and to cope with legacy. Objects and relations may have children to express different flavours and characteristics.

Despite of the exposed, to define criteria, logical formalization is required once facts and rules become part of the model. The logical paradigm uses Boolean[14] algebra to reason. Additional mathematical constructs give a more expressive power to differentiate concepts from each other.

Recommendation WP6_RC_F6: To design a conceptual model the object oriented paradigm SHOULD be used.

Recommendation WP6_RC_F7: A logical formalization SHOULD be used to define criteria.

---

[14] Boolean algebra, as developed in 1854 by George Boole in his book 'An Investigation of the Laws of Thought' is a variant of ordinary elementary algebra differing in its values, operations, and laws. Instead of the usual algebra of numbers, Boolean algebra is the algebra of truth values 0 and 1, or equivalently of subsets of a given set. The operations are usually taken to be conjunction $\wedge$, disjunction $\curlyvee$, and negation $\neg$, with constants 0 and 1. And the laws are definable as those equations that hold for all values of their variables, for example $x \curlyvee (y \curlyvee x) = x$. Applications include mathematical logic, digital logic, computer programming, set theory, and statistics. Source: http://en.wikipedia.org/wiki/Boolean_algebra

6.3. Meta constructs

Concepts and metadata are terminologies that are closely related. Metadata consists of data describing data. Whereas the concept is the definition for an entity, metadata are the real instances that serve different purposes. Metadata operationalize the concept for use in business processes.

The definition that an author is a person who has written a book refers to the entities author and book. A book you can buy is a real occurrence of the concept book. The information presented on the cover of the book and its ISBN-number are all metadata.

The following purposes are relevant:
- Provenance
- Supportive constructs for modelling
- To organize resources
- To organize content
- Improve automatic processing for documents.
- Additional data for work flow management
- Etc…

For the conceptual model additional attributes are necessary for provenance, version control and multilingual addendums.

6.4. The methodology

The methodology describes how uses cases can and should develop the semantics needed to support the business process executing the legal and judicial procedures. WP6 proposes a use case centric approach with an eye for reusability.

In real life this means a process to identify, retrieve, select or create concepts. The following procedure is followed:
1. Check the legislation. In case of a regulation, data definitions described in that regulation should be used. If data definitions have been described for directives than these data definitions must be used;
2. Check repository of data definitions for available data definitions with high acceptation/adaptation in the community;
3. Third and as the least preferred option, develop specific data definitions. The development preferably, but surely the result of definitions must be checked by institutions like for example SEMIC and ISA;
4. The repository will be made available and accessible for legal advisors and legislation writers in order to promote reuse of data definitions in EU-legislation, thereby enabling interoperability within the EU.

The information analyst and the ontologist or data analyst work together in close relation. The information analyst analyses the business process of the use case. The ontologist is an expert in different domains and keeps the general overview and knows how to relate those domains and to define the concepts.

The steps for them to follow are:
- Analyse the business process
- Identify existing concepts
- Define new concepts
- Place new concepts in context
- Create relations and connect the concepts
- Define rules to express criteria
- Do a consistency and quality check
- Identify concepts that become part of the logical model

When a use case centric approach is followed, the conceptual model is part of a three stage methodology. The conceptual model serves as input for the logical model.

The use of multilingual constructs is kept to a minimum in the conceptual world. They are meant for specific usage only. General metadata and criteria are part of the conceptual model and in one language only.

The paradigm is object oriented. Inheritance is a construct to specify parent child relationships. Specializations that have additional attributes become a child. Role modelling is kept to a minimum in the conceptual model. In the logical layer the building blocks are business interface and role specific. To incorporate legacy inheritance, it is useful to specify a broader or more specialized term. E.g. a person subject of law may be a suspect or a criminal; a drugs dealer is a type of criminal; in case of the criminal becoming offender, multiple inheritances keeps the structure intact.

7. The logical layer

7.1. A logical specification

The logical model contains specifications of constructs and components that become part of an electronic message, web service or repository.

The conceptual world gives the foundation for a common language. The concepts and their relations are part of this common vocabulary. To be semantic interoperable a further specification of data components is necessary.

The logical model is the layer where concepts turn into components. The concept and its definition are taken from the conceptual model and become a generic template for further usage. The generic template may be the logic counterpart of the concept natural person including its attributes and associations with other concepts. For a specific message to be sent and received belonging to a business interoperability interface a copy is made. This copy is to be fine tuned to meet the requirements set for the business interoperability interface. The natural person becomes a lawyer, judge or criminal with justifications made essential for the interface.

In contrast with the conceptual model the logical model incorporates supportive constructs. For instance a textual writing guides the reader in interpreting the data. Writing begins with a short introduction to the topic and gradually, moving forward from chapter to chapter, the reader consumes the contents and acquires information. The title, chapters, page numbers, index and appendix give the reader support in finding the right section. The same holds for the logical model where cardinality, options, footer, header, time stamp, references to other messages etc are important supportive constructs for a message.

The supportive constructs provide additional processing power to be interoperable. To model knowledge is a first step to be semantic interoperable. Both machine and human acquire knowledge to give a meaningful explanation for data input in the first place. Modelling knowledge is part of the conceptual world. The conceptual models give further context sensitive input. Common rules and structures on the logical level make machine processing possible.

Although, for machine processing a less fluid way to present the data is sufficient, for interpreting a document place holders are common to support a more flexible and context sensitive structure. The place holders guide a software program step by step to consume the document and give it an interpretation.

Different types of interaction lead to different place holders for data. Media that contain data are:

- Documents
- Web forms
- Message
- Container
- Data repositories

All these five types posses metadata and data. The structure and syntax is part of the logical specification. The specification itself is technology neutral.

Requirement WP6_RQ_NF2: The logical specification SHOULD be technology neutral

7.2. Meta constructs

Concepts are a good starting point to identify the data elements. In the logical specification more attention is given to data type specification and attribute selection.

The following constructs are present:

- Constructs to fine tune a concept
  - Reference both the generic template and a specialized version
  - Select the attributes to be part of a building block
  - Select a building block to be part of a large whole e.g. message or document
- Data types
  - Primitive, composite data types plus additional domain, range restrictions
- Code lists
  - Code plus definition

Code lists have a great semantic value in a European context. A code list is a combination of codes and their definitions. An update for both code and definition may remain independent. The definition is preferably in English to ease the process of reaching an agreement. The code, and not the definition, is leading. A good example for code lists is offered by ECRIS, the register for criminal records.

7.3. Orientation

A logical specification may lead to different implementations. The following three orientations influence the way we specify:

- Web service
- Message
- Repository

For these orientations, general and business interface specific specifications have to be in place. The Web service processes data through remote procedures. A specification for the data type and the procedure is to be made in e.g. the WSDL standard. The core of the message specification is formed by the structure, syntax and occurring data elements. The logical building blocks for the repository are schema, table, record and view.

For each orientation we have to know what type of data is to be sent, received or stored and in what format. The data to be sent may be the same, independent of the orientation chosen. A message or e-Service are in close relation. An e-Service may handle messages or posses functions to handle data elements (e.g. the web service is a special form of a e-Service) or a combination of both. The data present in the message may be fetched by a remote function or a combination of functions.

A logical specification may contain one or more of the following elements:
- The general document structure
- The general message structure
- The syntax to specify complex structures
- The syntax to specify functions that operate on the data
- The semantic rules
- Metadata to identify and classify a document, message or container
- Metadata to organize content for further processing
- Metadata to specify the internal structure for a repository

7.3.1. Document

Documents serve different purposes. For e-CODEX a document relates to a legal and judicial case. A document may have to meet different criteria. The authenticity is of high importance. WP4 'e-Identity' carries responsibility for the mechanisms to proof the identity of the author and authenticity of the document. Where a digital signature, watermark and time stamp are part of the document construct, rules when and how to use them are to be in place. A template gives guidance on how to write a document.

A document is unique and thereby owns an identity. The document is to be used in context, part of a procedure or direct communication. For direct communication in a letter, it is common to enter both information for the addressee and recipient in combination with the subject. For electronic communication, a unique reference for the document and so being able to place the document in context is a good practice. Documents are not subject of design in WP6. However it is important to know their nature and their unique reference.

> *Requirement WP6_RQ_F9:* For a document a unique identifier and provisioning information SHOULD be in place.

7.3.2. Web forms

Where an electronic message is for system–to-system integration, a web form is for human interaction. A decision making and guidance mechanism gives the user additional information in a flexible manner. The decision path is part of the conceptual model where logic is taken into account. To model the form, the data elements, the fields and their behaviour is an additional step that takes place on the logical layer.

A web form is for user input, and therefore placed on a website (or even available at a client server architecture) of a specific (national or EU) legal entity. e-CODEX cannot prescribe the technical format of these forms, but only the output of the contents. The result of the form has to be XML. Whether it is produced by a PDF-form, an XHTML form of a java-applet is out of scope.

> *Requirement WP6_RQ_F10:* A web form SHOULD deliver structured output or input in XML format for further processing

7.4. Information exchange agreement

A business interface is an agreement between Member States. A technical formalization on the logical level takes place. The interaction and the data require a further clarification. The business model describes the interactions. The logical model connects the data. Data may be sent in different forms and media.

The business interface specification contains one or more of the elements:
- A data element
- A business document, message or form
- A function

Step-by-step the elements are developed. The conceptual model gives essential information and acts as a template. The elements are built on basis of these templates. Selecting data types, code list and essential properties is done in logical model stage. A template is to be used more than once in different contexts. A business interface takes both the national and the supra-national perspective into account. Whereas the conceptual model identifies concepts and a direct translation of words, the logical level creates equivalent elements ready for use in daily communication. E.g. a bilingual dictionary compares words, a tourist phrase book gives common phrases in two languages that are ready for use in daily live. The information exchange agreement serves as a tourist phrase book: ready to use!

8.   The technical implementation

8.1.   The implementation, a technical perspective

The final stage for deliverance of a semantic building block is the technical implementation. The technical implementation uses the logical model for deriving the technical equivalent. Machine specific and software specific requirements are part of the technical specification.

8.2.   Character set

Whether XML Schema is generated or WSDL a character set should be present. UTF-8 is the standard for character encoding.

> *Requirement WP6_RQ_F11:* Character sets SHOULD be in Unicode UTF-8 format

8.3.   Referencing

A Unique Reference Identifier defines a namespace and a local name to reference resources (e.g. concepts, words and documents). A namespace shows the location and may point to an internet domain to ensure uniqueness. The local name is the resource that has a meaning in a certain context. The namespace gives direction which context is meant.

The namespace contains words separated by a period '.'. A namespace cannot have more than one meaning to be practical. It is a structure to separate homonym identifiers and to identify their domain.

URI come in different flavours.
- URNs are used for identification,
- URCs for including meta-information.
- URLs for locating or finding resources.

To reference a book both a URN and URL have their value. A book owns an ISBN (≈URN) for its identification and a URL to download the eBook version.

One can classify URIs as locators (URLs), or as names (URNs), or as both. A Uniform Resource Name (URN) is a name to identify a resource like a person's name identifies a human, while a Uniform Resource Locator (URL) is the reference to the location of the resource like a person's street address references the location a person lives. In other words: the URN defines an item's identity, while the URL provides a method for finding it.

The Functional Requirements for Uniform Resource Names are described in RFC 1737. The URNs are part of a larger Internet information architecture which is composed of URNs, Uniform Resource Characteristics (URCs), and Uniform Resource Locators (URLs). Each of them plays a specific role.

8.4. XML

To structure an electronic message or file, the standard XML gives a presentation independent structure. The character encoding is referenced in a XML document and should be UTF-8. The resources present in the document have a URI.

The style sheet is complementary to the XML standard. This format is suitable for transformation purposes.

8.5. Naming and Design rules

XML is a Meta format to define place holders. The syntax for a XML document is given for instance by a XML Schema. The name and the occurrences of the place holders occur in the logical model. Appliance of naming and design rules results in a XML Schema. The XML Schema is the physical implementation of the logical specification.

Naming and design rules prescribe a clear pattern how to name individual elements. A design pattern how to order the individual elements, their references and definitions complement those rules.

> *Requirement WP6_RQ_F12:* The naming of individual data elements SHOULD be compliant with naming and design rules.

9. Trans-border e-Government services

9.1. e-Service

In paragraph 4.4 three different types of interaction are distinguished, together with four levels of sophistication for communication.
The MS or the e-CODEX project should provide e-Services to enable users (citizen or business) to communicate on a level of their choice.

By an e-service, in the context of e-CODEX, we mean a service that resides on a website on the internet, that is freely accessible to all intended users, and that performs a well-defined task.

A website or a portal should feature a list of all e-Services available. Each e-service has several ways of implementing the service. The documents needed for interactions of the business-to-government or citizen-to-government types must be made available, listed under the appropriate e-Service. Additional services are needed to sign a document, to attach metadata concerning the sender, type of message, date of creation etc, and to attach an unlimited amount of unstructured documents.

EIF provides a list of requirements for e-Services in the paragraph 'Business requirements for e-Government services'. In addition, the following requirements can be identified:

Requirement WP6_RQ_NF3: The Commission or MS's should, preferably by the e-Justice portal, provide the necessary services to enable a user to create a document using the document standards.

Requirement WP6_RQ_F13: The available services SHOULD be accessible to all intended parties.

9.2. Business process

The activities for business process analysis are in WP3, 5 and 7. The recommendations will be addressed by these WPs. The Business Process is the sequence of activities which is performed by an organisation or multiple organisations in order to achieve their business goals. Regardless of any support by ICT means, the business process can be performed. With the support of ICT means, the execution of these actions is simpler, faster and less exposed to human error.

Looking at the different levels of interoperability according to the EIF interoperability levels, it can be stated that the business process is in the layer of 'organisational interoperability'. It defines how different organisations are aligned to achieve common goals. Also, looking at the EIF levels, the business process is formed by national and European regulations which live in the 'Legal Interoperability' layer. Finally, the Business Process is the input for the 'semantic interoperability' level.

In general, the goals of business process supported by the e-CODEX project are not so much organisational goals, but goals of the European society, focused on legal and judicial matters. Organisations will need to cooperate and consider goals of the society more important than their own organisational goals. Their combined knowledge, workflow and information are needed in order to best serve citizens, businesses and governments.

In the context of e-CODEX, a Business Process defines the interactions between two or more partners in order to achieve a common information status regarding a certain business process. Generally these interactions follow a predefined sequence which the collaborating partners have agreed upon beforehand. It is only logical that a seller cannot confirm an order if the buyer has not placed one. As shown in the figure (figure 1: Business Transaction Example) below.
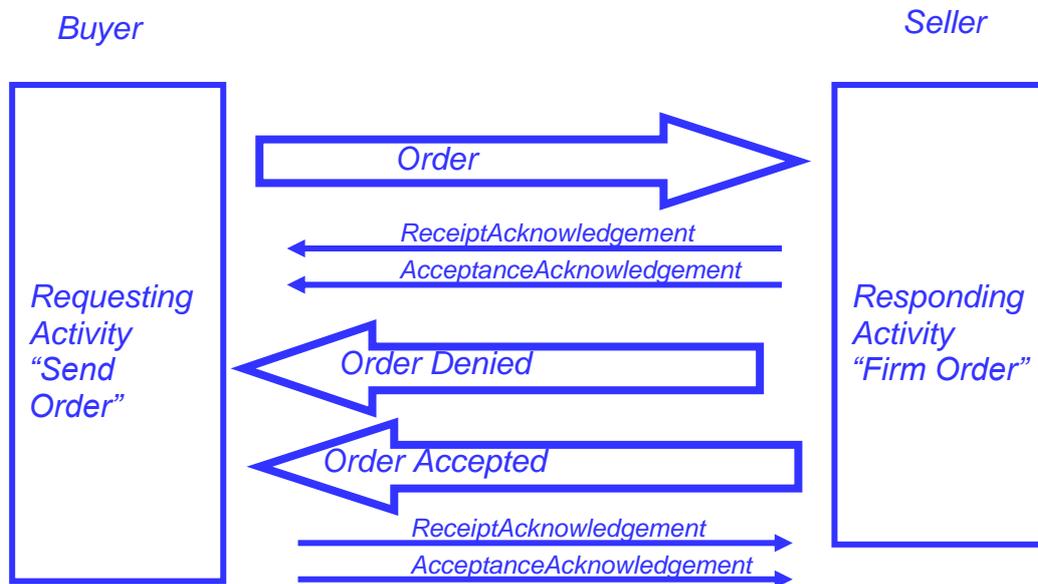
Figure 1: Business Transaction Example

On a higher level: the seller will not deliver goods before an order is placed and a buyer will not pay before an invoice is received. These higher lever examples illustrate the distinction between an interaction and a transaction; a transaction describes a coherent set of information exchange. i.e. 'Ordering': Order --> Order confirmation, whereas an interaction describes a coherent set of transactions: i.e. 'Buying': Ordering, Invoicing, Paying, Delivering.

The lowest level in a business process is the exchange of a business document itself. The actual information exchange is referred to as the exchange of a business document. In the figure above examples of business documents are: 'Order', 'Order Denied' and 'Order Accepted'. Once collaborating partners have established an understanding of what happens (interactions and transactions) when (the sequence of interactions and transactions), they will know what messages (business documents) to send and receive. Now all they need to establish is the content of the business documents. What is in the business documents is derived from the answer to 'what information does one need to perform this particular task?'
The technical equivalent for a business document is a XML document.

By defining the information need and tagging these information elements as metadata elements, we have entered the level of semantic interoperability. The hierarchy of a business process is:
- Interactions
- Transactions
- Business documents

9.3.   Infrastructure

Through the infrastructural layer, collaborating partners are interconnected. I.e.: their respective infrastructures are interconnected in order to exchange data derived from their internal business processes, supported by business applications.

Ideally, technical interoperability is achieved by interconnecting organisations, rather than interconnecting applications. In this way the independency of organisational infrastructure is maintained. At the 'front door' of each organisation an interface has to be put in place to connect to other infrastructures with other specifications. If an organisation has to interconnect to multiple other organisations, with possible multiple other infrastructure specifications, it has to avoid creating multiple interfaces. Therefore a common mutual infrastructure has to be used, to which each organisation can connect with just one interface. This approach is in line with the star topology architecture.

Although this mutual infrastructure is in the responsibility of WP5, a short statement on this topic is in place. For routing and addressing details, on which the mutual infrastructure depends, WP6 has to semantic support for addressing and routing information as well as definitions for error handling.

Logical addresses for participating organisations in each member state, naming and design rules for interactions, transactions and business documents are needed for cross border routing of messages and internal routing to applications within the respective organisations. These attributes will be put in the header of messages, and the business document itself will have the payload shown in the figure (figure 2: Message Structure) below.
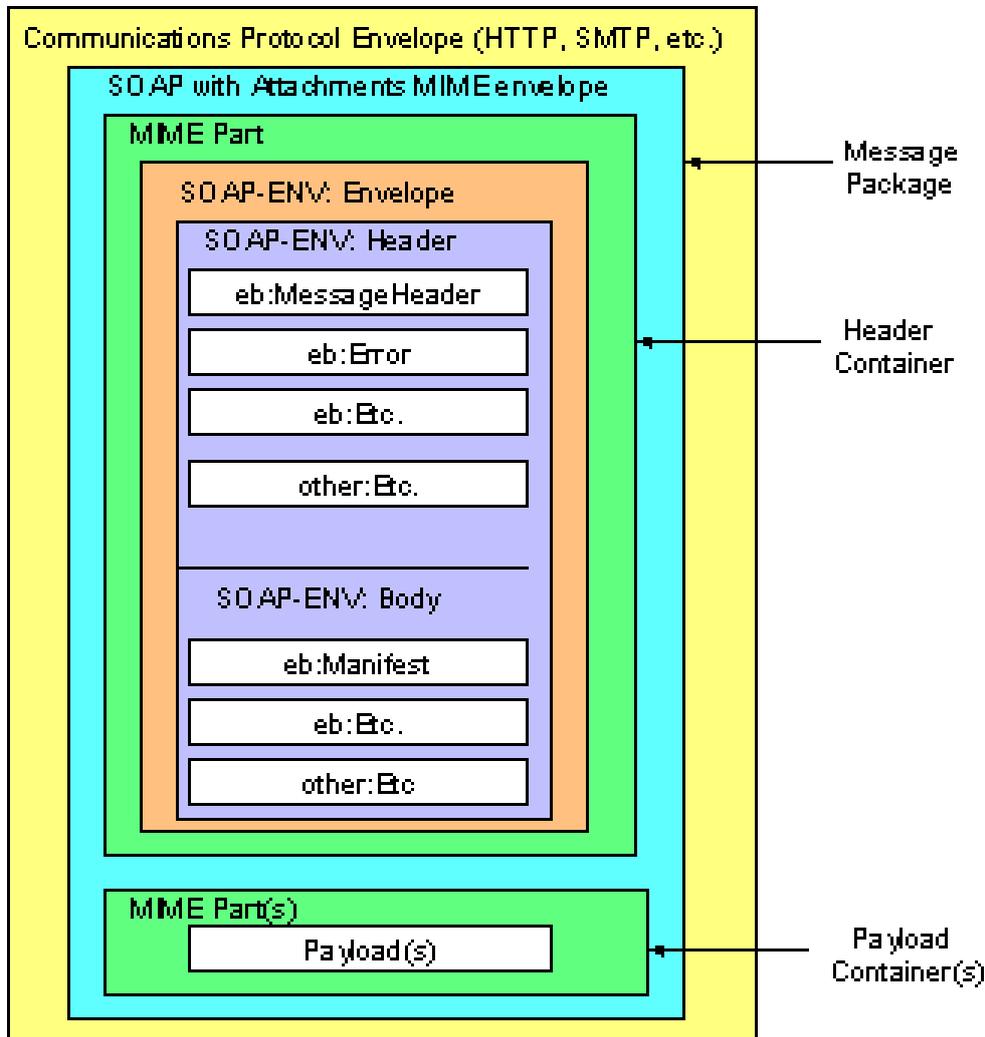
Figure 2: Message Structure

9.4. e-Service integration

Figure 3 Service Loose Coupling Interoperability Layers reflects the integration of the e-Service for e-CODEX. The e-Service consists of 5 components which together will support all procedures. These components are:

- Business Process of collaboration partners bound together
- Semantic support
- Business applications of collaborating partners
- Internal infrastructures of collaborating partners
- Mutual infrastructure to interconnect collaborating partners.
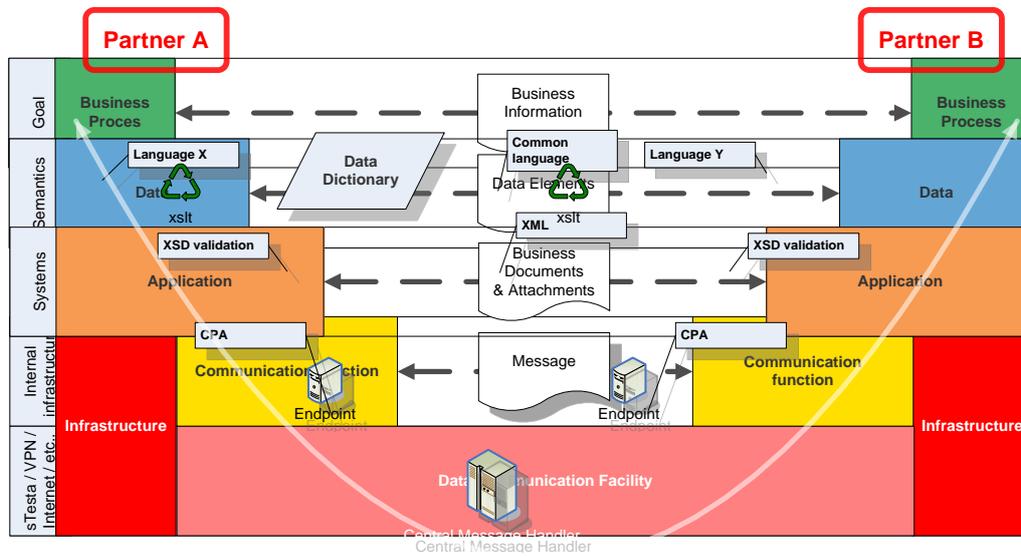
Figure 3: Service Loose Coupling Interoperability Layers

Only if these components are integrated in optima forma, the e-Service will be effective. Since the key for a successful e-Service is a smooth execution of a business process collaboration, which in turn is dependent of the alignment of information status, the semantic aspect of e-Services is evident. To ensure that information is exchanged reliable, safe and fast by the infrastructures of the collaborating partners have to be aligned through a common infrastructure.

10. e-Service: human interaction

10.1. Publication

Publication is the distribution of content to the public or enterprises. Instructions about the usage of the e-Services, legal advice and reference to legal procedures have to public accessible.

Recommendation WP6_RC_NF1: Publication SHOULD been put under the control of a legal body.

10.2. Form

To gather information concerning a judicial or legal procedure a standard form is technology neutral by nature. Electronic communication requires one or more documents standards to be eligible for processing. A form may be integrated with an e-Service or may be a separate document.

It is recommended that forms features checks for correct input to limit mistakes and to enhance the probability the forms can be processed. Completed forms should result in XML-messages, with a digital signature if required.

The process in which or the way the user is guided in using the service is out of scope for WP6. WP6 specifies the output of the service, not the process for creating a document, nor the user interface.

10.3. Intelligent form

An intelligent form gives guidance to the user to complete a form. A decision path makes clear which options are open to the user. A decision made may lead to different follow up actions. In providing only those options relevant for the present context mistakes are prevented.

The conceptual model and logical model part of the methodology to construct XML Schema give input for the constructing process for an intelligent form. The conceptual model may incorporate axioms, facts, relations and rules essential for creating a decision path.

Designing an intelligent form is a specialist's work. The result of an intelligent form should be a user validated structured document, meeting the requirements of an agreed upon document standard. Designing an intelligent form is out of scope of WP6.

11. e-Service: system to system

System to system integration is the first priority for e-CODEX and WP6. The methodology with three layers, the conceptual model, the logical model and technical implementation give step by step input to the message specification. The specification describes what data may be present and how it is to be combined in the message. The technology in use is XML. A message is a XML document that is compliant with the specification, the XML Schema, present for the type of message.

11.1. Service functions

With system to system integration the user is not directly involved in the details how the message is constructed and how the delivery is made. The creation of the message is done by an e-Service. The eService displays a user interface where the user may input his or her data.

Concerning this fact the following recommendations for the e-Service are made.

Recommendation WP6_RC_F8: An e-Service SHOULD been made known to users and users SHOULD been made aware of the benefits of using the service.

Recommendation WP6_RC_F9: An e-Service SHOULD be user-centred.

Recommendation WP6_RC_F10: An e-Service SHOULD add value and posses a clear function part of business process.

The e-Service is responsible for the creation of the XML document compliant with a predefined XML Schema. The e-Service possesses the necessary functionality to retrieve data from an electronic form and transform it in a XML document ready to be sent.

Requirement WP6_RQ_F14: An e-Service SHOULD be responsible for the XML document creation process.

Requirement WP6_RQ_F15: The details of creation a XML document SHOULD be hidden for the user.

11.2. Messaging

An overall European dictionary contains the elements that are generic in nature and may be in use for more than one use case. The use case specific one contains those elements necessary for the judicial or legal procedure part of the use case. For reuse purpose and to be transparent, public access to a repository containing the dictionaries should be present. The repository can and will host multiple dictionaries. The naming and design rules will set a standard to tell users of any kind about the source of elements. A not yet identified legal body will be responsible for maintaining the repository.

The message specifications must be conformant to a set of naming and design rules, to enable a receiving system to validate the message. To sustain the naming and design rule standard a standardisation forum must be present.

Requirement WP6_RQ_F16: A message specification SHOULD contain only those elements/components part of the European dictionary and use case specific one.

Requirement WP6_RQ_NF3: For public access a repository for all dictionaries SHOULD be present.

Requirement WP6_RQ_NF4: The repository SHOULD be maintained by a legal body.

Requirement WP6_RQ_NF5: The naming and design rules SHOULD be maintained by a standardisation forum.

Partners standardise on the logical specification of the XML message. Each partner may have additional business rules to be met. Those rules may imply additional restrictions for the XML message to be processed. Partners therefore should make information exchange agreements for each interaction for each business process.

Recommendation WP6_RC_F11: Support for verifying a XML document to be compliant to a set of additional business rules SHOULD be present.

For the XML message more than one logical equivalent may be present to support the national specific requirements to be met. For further processing the XML message may have to be transformed.

Requirement WP6_RQ_F17: Transformation of logical building blocks SHOULD be possible to sustain national legacy.

12.   Conclusion

WP6 document standard and semantics gives a general set of requirements. This set is the foundation for more use case specific additions to come. All these requirements are in line with the EIF recommendations. With a three layer architectural framework the first contours for semantic interoperability within eCodex are made.

WP6 focus is on system to system integration. Structured electronic messages and the data elements part of these messages are the major topics. The standard to use for electronic messages is XML.

To construct the right message specification a methodology must be in place. WP6 incorportates the object oriented paradigm, three layer architecture and standardizes on naming and design rules.

To deliver semantic components for the use cases is the first step in reuse. A repository with generic components must be available to broaden the possibilities for further initiatives for cross border information exchange.

The method of work and the requirements presented in this document contribute to the sustainability of solutions developed in e-CODEX.

I. Appendix

In this Appendix a high level overview is given of SPOCS and PEPPOL as two important LSP's. An extensive analysis is of these LSP's as well as of national solutions and others is given in D6.2, "List of standards, reusable assets to be used and missing building blocks".

A.1 PEPPOL

A.1.1 Global overview

PEPPOL (European Public Procurement Online) works on the realisation of an EU wide public eProcurement. National systems connect to form a unified European market.

PEPPOL presents for the semantic interoperability layer a methodology, standards for a dossier and services. PEPPOL focuses on streamlining the process for data exchange, the information for enterprise to qualify and providing proof for this qualification.

A key enabler for semantic interoperability is the Virtual Company (VCD). To share documents that relate to each other a VCD gives a solid structure. An e-Signature proofs the authenticity, for the dossier and the individual documents. Metadata provides additional information, the subject, the source, the purpose, the criteria to be met and the relation with other resources. Both for the document and the dossier metadata is present. For Government to government, government to public and government to business a VCD ensures the sustainability and longevity in a European context.

A.1.2 Interaction form

The interaction form classifies as an intelligent manual process with reasoning support and machine translation. Tools show the documents to the human operator.

A.1.3 Possibilities for adoption

The VCD might be an answer for interoperability on legal and judicial matters. For legislation the European and nation legislation prescribe criteria to be met. Formalising these criteria and verifying them during the process is a huge step forward. The VCD system supports mutual recognition of evidence from different origin.

The logic to formalize a criterion knows a broader usage. In essence concepts, axioms and rules as a whole define the set of criteria. Information extraction and reasoning give a faster way to identify crucial information, especially in a multilingual environment.

A.1.4 Standards and methodology

The process to formalize criteria starts with European legislation. For PEPPOL EU directives form the foundation for the national interpretation. The evidence is judged on bases of the criteria. The criteria ensure consistency and equal treatment.
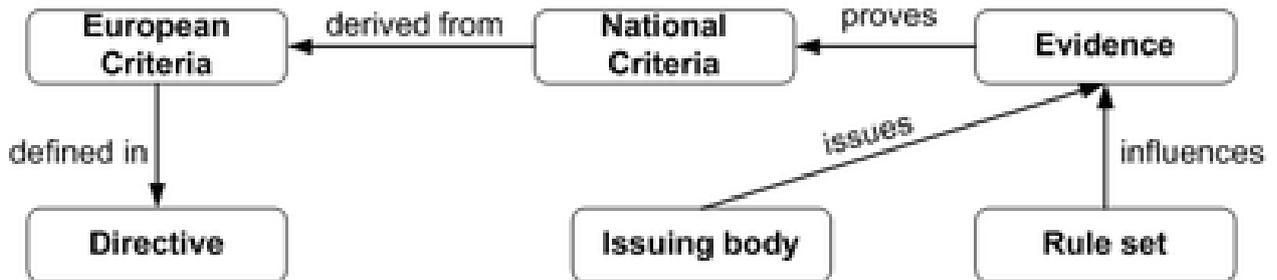
Figure 4 Semantic interoperability model of the European VCD service
source www.peppol.eu

Mapping criteria is a job for legal experts and ontologists. A template is used in the first stage to acquire the rule sets. The mapping is made more explicit by logical formalisation. All three subjects criteria, evidence, tenderer are placed in separate ontology. A fourth one enlists concepts general in nature. The methodology is model driven nature. A fifth ontology contains input for the reasoning process. PEPPOL uses the W3C standard OWL.

A four stage model for VCD gives an easy entry level for European interoperability. The semantic interoperability strengthens with the advance in maturity level. The first stage ensures recognition of the certificates and qualification documents. The second stage defines metadata to guide the electronic process of compiling, enclosing and submitting documents.
The third stage introduces a modern technique to prove evidence with machine support. The fourth stage gives additional services to submit a document and streamlines the process further with support to reference, up and downloading the document.

A.2 SPOCS

A.2.1 Global overview

SPOCS is a European Union initiative to provide centric oriented services where one face is shown to the public. The Member States communicate with their citizens through a single point of contact.

An example:
For people who start a European career you have to attest your credentials. Your home town where you live, your school, your previous job, your curriculum vitae are information you have to prove to be genuine. Settling in the European Union means you have to be in contact with your mother land.

The single point of contact gives light to the principle no wrong door. People who direct to their government with a request for a service get access with a simple link to a web site.
A single point contact is responsible for providing services concerning procedures to follow. Information on the applicable procedure is present and people are untitled to get information in their own language. Interaction with other Member States is fewer burdens for the citizen of state for he or she can communicate through the same single point of contact.

SPOCS knows the following core functionality:
1. Content Syndication
2. Reliable message exchange
3. Electronic signing documents
4. Verification of authenticity
5. A digital dossier with all files part of a procedure
6. Digital safe to store a digital dossier, statements, permits and diplomas
7. Deliverance and response to procedural request.

A.2.2 Interaction Form

SPOCS is a show case for human computer interaction where system to system communication is on the back ground. A citizen is presented a single point of contact and easy to use interface.

A.2.3 Possibilities for adoption

The ideal content syndication creates a constellation where the public services are redirected through a single point of contact.

The digital safe and the electronic dossier OCD are to be taken under consideration. The OCD format is an addition to the VCD standard. Two containers a PDF version and a ZIP file are present. SPOCS gives an example how to strengthen a loosely coupled multi lingual federation. Metadata for both the content and the dossier in relation to procedures is the linking pin.

A.2.4 Standards and methodology

To format documents structure or structure less comes with pros and cons. For a dossier documents share a common back ground. The relation they have is to be expressed with metadata. SPOCS works with the OCD container format and works with the two VCD metadata layers for the overall dossier structure and the content for a single document.

SPOCS gives a good addition to PEPPOL and in combination gives a richer set of services to work with documents and containers. SPOCS gives an easy entry level for introducing new services. Services to generate a container, support for creation of metadata, importing the payload and signing both the OCD and the individual documents and metadata are in place.

A combination of light computer aided support and manual action for the first stage gives flexibility where the conditions to be met are uncertain. With the growth of understanding of the process to deal with the higher levels of maturity that are part of PEPPOL come into play.
SPOCS concept of content syndication links the national services for each Member States where the business may evolve through time in a organic manner.

A.3 Eurojust

A.3.1 Global overview

Eurojust is a judicial cooperation body created to help provide safety within an area of freedom, security and justice. Eurojust was set up by the Council of the European Union in February 2002 ([Council Decision 2002/187/JHA](#)) to improve the fight against serious crime by facilitating the optimal co-ordination of action for investigations and prosecutions covering the territory of more than one Member State with full respect for fundamental rights and freedoms. Eurojust is composed of 27 National Members, one from each EU Member State. These are senior and experienced judges, prosecutors, or police officers of equivalent competence, who together form the College of Eurojust.) Eurojust has adopted a large part of the US NIEM framework, and more specifically the Global Justice XML Data Model. The GJDXM has been created for the US judicial system, and as such not suitable for complete adoption in a European setting. So, the model was adapted in part to suit the European needs.

Eurojust is the only party in Europe to have adopted NIEM. Although several parties have stated their interest in the standard, none of these have adopted NIEM.

The NIEM standard is to be studied in depth in WP6, to evaluate its usability for the e-CODEX project.

A.4 SEMIC

A.4.1 Global overview

The SEMIC website states that 'SEMIC.EU is a participatory platform and a service by the European Commission that supports the sharing of assets of interoperability to be used in public administration and e-Government.

The Semantic Interoperability Centre Europe (SEMIC) is a service provided by the European Commission. It was initiated in the framework of the programme 'Interoperable Delivery of European e-Government Services to public Administrations, Businesses and Citizens' (IDABC). SEMIC features a repository of assets, contributed by the member states, which can be downloaded and re-used, to promote semantic interoperability.

SEMIC has recently published a component Natural Person, which can be used by all member states. This component has been defined in collaboration with data modellers from several member states.

In general the components in the asset repository are not all of the same quality or completeness. The components in the repository need to be further investigated, to assess their reusability for the e-CODEX project.

II. Appendix Requirements

In this Appendix the different requirements and recommendations described in the different chapters of this document are summarised and categorised.

Identifier: WP6_RQ_F1	Name of Requirement: EIF
Description	e-CODEX SHOULD incorporate the EIF recommendations and SHOULD refine them for legal and judicial application.
Comments	The European Interoperability Framework is the foundation for interoperability in European. Essential recommendations how to be interoperable are present. WP6 takes those to be mandatory.
Reason / Benefit	Reuse of broad accepted requirements

Identifier: WP6_RC_F1	Name of Recommendation: Information architecture
Description	The information architecture for WP6 SHOULD be formed by a three layer methodology that identifies the conceptual model, logical model and technical implementation.
Comments	An architecture framework for the semantic interoperability layer (according to EIF) and a methodology how to design the semantic components is essential. A clear separation in three layers with a different purpose simplifies complexity issues. The conceptual model for identification and relating entities. The logical model to build components ready for in use in cross border message exchange. The technical layer the implementation of message specifications or web service specifications.
Reason / Benefit	Abstraction and dividing complexity result in higher quality.

Identifier: WP6_RQ_F2	Name of Requirement: European core model
Description	An European core model SHOULD be present for e-CODEX to link sector and use case specific models.
Comments	Models with semantic components share a

	central core model with components common to all of them. This to be able to transform and enhance reusability and good design.
Reason / Benefit	Enabling transformation, reusability, good design and to provide a common foundation to build on.

Identifier: WP_RQ_F3	Name of Requirement: Modelling data exchange
Description	Data elements for data exchange purpose SHOULD be defined on the logical model layer.
Comments	The logical model defines the semantic components ready for use in direct communication.
Reason / Benefit	A clear abstraction where the specification and technical counterpart are separated gives higher quality and is free of technical driven aspects.

Identifier: WP_RQ_F4	Name of Requirement: Mapping
Description	Member States SHOULD map their national data models with the European core model.
Comments	MS are free in developing their own national data models. To be semantic interoperable national semantic components have their European counter part and vice versa.
Reason / Benefit	Subsidiarity and mapping are essential to be semantic interoperable.

Identifier: WP6_RQ_F5	Name of Requirement: Preservation
Description	The original format of a document SHOULD remain intact during transportation.
Comments	Preserving information and preserving the carrier during transport is a perquisite to ensure the integrity.
Reason / Benefit	Integrity of data

Identifier: WP6_RQ_NF1	Name of Requirement: Digital archive
Description	European and national guidelines SHOULD be leading concerning digital archiving legal cases.
Comments	To be semantic interoperable the original

	information must be retrievable. Digital archiving is the solution for persistent storage.
Reason / Benefit	Standardisation, persistency and accessibility of data.

Identifier: WP6_RQ_F6	**Name of Requirement: Identifier**
Description	All meaningful objects which are exchanged SHOULD have unique and persistent identifiers for later reference.
Comments	To retrieve information one must be able to identify the object containing the information.
Reason / Benefit	Accessibility, recall and precision in retrieving information.

Identifier: WP6_RQ_F7	**Name of Requirement: Identification**
Description	Where identification systematics for legal information are available they MUST be used.
Comments	Standardisation how to identify legal information makes the legal information accessible.
Reason / Benefit	Standardisation and interoperability

Identifier: WP6_RC_F2	**Name of Recommendation: Dublin Core**
Description	The Dublin Core SHOULD be in use to describe resources.
Comments	The Dublin Core contains a prescribed set of elements to describe resources.
Reason / Benefit	Standardisation and interoperability

Identifier: WP6_RC_F3	**Name of Recommendation: Inter Lingua**
Description	An inter lingua SHOULD be present to develop a conceptual model.
Comments	A conceptual for a broad audience is best to be designed in a language everybody understands.
Reason / Benefit	Standardisation and interoperability

Identifier: WP_RC_F4	**Name of Recommendation: The conceptual model**

Description	A conceptual model SHOULD serve the following purposes:
	- Identifying, defining and relating concepts - Concepts lead to general metadata elements for classifying documents, messages and web pages. - Modelling metadata elements to uniquely identify content and processing purpose - Building a rule set for criteria to be met Input for the logical building blocks which have additional attributes.
Comments	For semantic interoperability purpose a conceptual model incorporates entities, relations, standard rules to infer additional relations. The entities are abstract concepts, with a supportive role, or real objects. To find existing entities or defining new ones is a modelling task which uses the conceptual model. With a conceptual model that holds expressive power with a rich set of entities and their relations representing a domain eases the task to classify information or to build semantic components. To classify information and to prescribe criteria to be met gives a firm grip on this information.
Reason / Benefit	Methodology

Identifier: WP_6_RC_F5	Name of Recommendation: Use case centric
Description	A use case centric approach SHOULD be followed. The conceptual model regarding the use case SHOULD posses concepts for this use case only .
Comments	WP6 is use case centric by nature. The conceptual model describes the use case.
Reason / Benefit	Non-redundancy

Identifier: WP6_RQ_F8	Name of Requirement: A technology neutral conceptual model
Description	The conceptual model SHOULD be technology

	neutral.
Comments	The conceptual model focuses on the information to be presented. Technology specific concepts are out of the question.
Reason / Benefit	Technology neutral

Identifier: WP6_RC_F6	Name of Recommendation: The object oriented paradigm
Description	To design a conceptual model the object oriented paradigm SHOULD be used.
Comments	The object oriented paradigm owns powerful constructs to structure objects and to reuse specific assets.
Reason / Benefit	Modularity, flexibility and quality of design

Identifier: WP6_RC_F7	Name of Recommendation: Criteria
Description	A logical formalization SHOULD be used to define criteria.
Comments	Criteria have to be formalised. The logical formalisation ensures a precise interpretation.
Reason / Benefit	Disambiguation and precision of the criterium.

Identifier: WP6_RQ_NF2	Name of Requirement: A technology neutral logical model
Description	The logical specification is SHOULD be technology neutral.
Comments	The logical specification of semantic components for cross border exchange makes the adoption of different technologies possible and makes interoperability more durable.
Reason / Benefit	Technology neutral

Identifier: WP6_RQ_F9	Name of Requirement: Provisioning
Description	For a document, a unique identifier and provisioning information is to SHOULD be in place.
Comments	The document owns a unique identifier and additional information of its origin and creator.
Reason / Benefit	Traceability of the origin and life cycle.

Identifier: WP6_RQ_F10	Name of Requirement: Structure data
Description	A web form SHOULD deliver structured output or input in XML format for further processing
Comments	To be interoperable, e-Services transfer data with a standardised structure.
Reason / Benefit	Standardisation and reliable processing

Identifier: WP6_RQ_F11	Name of Requirement: UTF-8
Description	Character sets SHOULD be in Unicode UTF-8 format.
Comments	UTF-8 is the standard for character encoding. An encoding is essential to interpret the characters consistently.
Reason / Benefit	Standardisation for system independent processing.

Identifier: WP6_RQ_F12	Name of Requirement: Naming and design rules
Description	The naming of individual data elements SHOULD be compliant with naming and design rules.
Comments	Naming and design rules ensure a consistent structure.
Reason / Benefit	Consistency

Identifier: WP6_RQ_NF3	Name of Requirement: Human interoperability
Description	The Commission or MS's SHOULD, preferably by the e-Justice portal, provide the necessary services to enable a user to create a document using the document standards.
Comments	For a user to transfer information, the information must be presented in the format in use. The user should have no burden with technological details.
Reason / Benefit	Openness, availability and user friendliness

Identifier: WP6_RQ_F13	Name of Requirement: Accessibility
Description	The available services SHOULD be accessible to all intended parties.

Comments	For a service to be used it should be accessible.
Reason / Benefit	Accessibility

Identifier: WP6_RC_NF1	**Name of Recommendation:** A legal body
Description	Publication SHOULD be put under the control of a legal body.
Comments	For quality and further development a party has to maintain the content. A legal body owns the right qualifications.
Reason / Benefit	Quality assurance and sustainability

Identifier: WP6_RC_F8	**Name of Recommendation: Service awareness**
Description	An eService SHOULD been made known to users and users SHOULD been made aware of the benefits of using the service.
Comments	For semantic interoperability and especially an e-Service it's essential everybody is able to know the benefits and purpose.
Reason / Benefit	Awareness of the possibilities for the public.

Identifier: WP6_RC_F9	**Name of Recommendation: User centric**
Description	An eService SHOULD be user-centred.
Comments	A human friendly e-Service is the foundation of semantic interoperability in Europe.
Reason / Benefit	Human friendliness

Identifier: WP6_RC_F10	**Name of Recommendation**: A clear out line
Description	An eServices SHOULD add value and posses a clear function part of business process.
Comments	
Reason / Benefit	Transparency of the functionality offered and prevention of duplication.

Identifier: WP6_RQ_F14	**Name of Requirement: XML creation process**
Description	An eService SHOULD be responsible for the XML document creation process.
Comments	To prevent mistakes the XML document is not to

	be created by man. In the context of WP6 an e-Service has to own functionality to process and build XML document.
Reason / Benefit	Completeness of essential functionality to be interoperable.

Identifier: WP6_RQ_F15	Name of Requirement: Hiding detail
Description	The details of creation a XML document SHOULD be hidden for the user.
Comments	The human is not to be bothered with technical details.
Reason / Benefit	User friendliness

Identifier: WP6_RQ_F16	Name of Requirement: Correct and complete
Description	A message specification SHOULD contain only those elements/components part of the European dictionary and use case specific one.
Comments	It must be clear which data elements to expect. Standardisation on data elements ensures reusability and a clear understanding.
Reason / Benefit	Completeness and correctness of the specification presented. Preventing redundancy or deviation.

Identifier: WP6_RQ_NF3	Name of Requirement: A metadata repository
Description	For public access a repository for all dictionaries SHOULD be present.
Comments	Public access promotes the data elements for further adoption.
Reason / Benefit	Public accessibility

Identifier: WP6_RQ_NF4	Name of Requirement: A legal body
Description	The repository SHOULD be maintained by a legal body.
Comments	For quality and further development a party has to maintain the content. A legal body owns the right qualifications.
Reason / Benefit	Quality assurance and sustainability

Identifier: WP6_RQ_NF5	Name of Requirement: Standardisation forum
Description	The naming and design rules SHOULD be maintained by a standardisation forum.
Comments	A standard becomes a standard when there is adoption and maintenance to ensure durability.
Reason / Benefit	Promotion and adoption of methodology. Ensuring stability and further development.

Identifier: WP6_RC_F11	Name of Recommendation: Business rules
Description	Support for verifying a XML document to be compliant to a set of additional business rules SHOULD be present.
Comments	The business requires mostly additional business rules for further processing.
Reason / Benefit	Consistency

Identifier: WP6_RQ_F17	Name of Requirement: Transformation
Description	Transformation of logical building blocks SHOULD be possible to sustain national legacy.
Comments	Semantic components are input for different systems. Different transformation of the same components must be possible to support those. A clear interpretation of the transformation ensures quality. Ambiguity has to be prevented when possible. To support different legacy systems is challenging.
Reason / Benefit	Interoperability for legacy systems

Epilogue

A first impression gives new fruit of thought. To define requirements and to agree on the first step may seem a huge undertaking. Semantics is a challenging journey where business cases may vary and have their own characteristics.

The art of drawing the path of evolution and to reach the same heights as the poem *I wrote her name in the sand* makes us move forward. It is great to see the European Commission starts new initiatives so the writing in the sand endures the test of time.

Semantics is a continual work in progress. Our second stage is to learn more from EU initiatives like PEPPOL, SPOCS, STORK, SEMIC and Eurojust.